www.embojournal.org

THE EMBO JOURNAL

Proofreading dynamics of a processive DNA polymerase

Borja Ibarra[1,7], Yann R Chemla[1,8], Sergey Plyasunov[1,9], Steven B Smith[1], José M Lázaro[2], Margarita Salas[2] and Carlos Bustamante[1,3,4,5,6,*]

[1]Department of Physics, University of California, Berkeley, CA, USA, [2]Centro de Biología Molecular 'Severo Ochoa' (CSIC-UAM), Madrid, Spain, [3]Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA, [4]Department of Chemistry, University of California, Berkeley, CA, USA, [5]Howard Hughes Medical Institute, University of California, Berkeley, CA, USA and [6]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Replicative DNA polymerases present an intrinsic proofreading activity during which the DNA primer chain is transferred between the polymerization and exonuclease sites of the protein. The dynamics of this primer transfer reaction during active polymerization remain poorly understood. Here we describe a single-molecule mechanical method to investigate the conformational dynamics of the intramolecular DNA primer transfer during the processive replicative activity of the Φ29 DNA polymerase and two of its mutants. We find that mechanical tension applied to a single polymerase–DNA complex promotes the intramolecular transfer of the primer in a similar way to the incorporation of a mismatched nucleotide. The primer transfer is achieved through two novel intermediates, one a tension-sensitive and functional polymerization conformation and a second non-active state that may work as a fidelity check point for the proofreading reaction.
The EMBO Journal (2009) **28,** 2794–2802. doi:10.1038/emboj.2009.219; Published online 6 August 2009
Subject Categories: genome stability & dynamics
Keywords: bacteriophage Phi29; DNA replication; molecular motors; optical tweezers; proofreading

Introduction

Replicative DNA polymerases are responsible for the faithful replication of DNA before cell division (Benkovic *et al*, 2001). They contain a polymerization (pol) active site that uses one strand of the DNA as a template and precursor deoxynucleoside triphosphates (nucleotides) that are added onto the 3′ end of the growing primer terminus to synthesize a new complementary (or primer) strand. The incorporation of a mismatch perturbs the specific interactions of the pol domain with the primer–template duplex, and specifically 6 out of all 12 possible mismatches also perturb the interactions with the template strand, which is displaced from the pol active site (Johnson and Beese, 2004). These structural perturbations stall replication and induce the transfer of the primer strand to the 3′–5′ exonuclease (exo) active site, where the erroneous nucleotide is hydrolysed from the chain (for a review see Kunkel and Bebenek, 2000). For some replicative DNA polymerases, the transfer reaction is carried out in an intermolecular manner between the pol active site of one molecule and the exo active site of a different polymerase molecule, which in some replication systems, such as T4, could be associated with the processivity clamp (Yang *et al*, 2004). After hydrolysis at the exo site, the trimmed primer strand is transferred back to the pol site processively (Strick and Knopf, 1998; Fidalgo da Silva and Reha-Krantz, 2007; Hogg *et al*, 2007). Instead, for some other replicative polymerase-DNA, such as T7 and Phi29, there is a bidirectional communication between the active sites and the transfer reaction is carried out intramolecularly (Donlin *et al*, 1991; De Vega *et al*, 1999). Interestingly, the pol and exo sites are typically separated by 30–40 Å (Freemont *et al*, 1988; Wang *et al*, 1997; Kamtekar *et al*, 2004), and the primer transfer reaction requires the opening of 3–4 bp of the primer–template structure for the primer strand to reach and be stabilized at the exo site, in a process that may involve several intermediates (Shamoo and Steitz, 1999; Hogg *et al*, 2004; Subuddhi *et al*, 2008).

Remarkably, during the processive DNA replication, a tight and fine-tuned coordination between the pol and exo cycles should exist to permit a productive but at the same time faithful replication. The dynamics of this coordinated action and the kinetic intermediates that are involved and their interconversion remain largely unsolved.

Here we report a single-molecule manipulation method to study the dynamics of the partitioning mechanism by applying varying tension to a processive single Φ29 polymerase–DNA complex. Φ29 DNA polymerase presents the highest processivity described so far for a DNA polymerase (⩾70 kb), without the aid of additional processivity factors (Blanco *et al*, 1989; Rodriguez *et al*, 2005) and as part of its proofreading mechanism it transfers the primer strand in an intramolecular manner between the pol and a built-in exo domain (De Vega *et al*, 1999). We show that the application of mechanical force to the template triggers the gradual intramolecular transfer of the primer between the active sites of the protein. To elucidate this transfer pathway, we compare the response to tension of the wild-type (wt) polymerase with that of two polymerase variants defective in different steps of the proofreading process: an exo-deficient (ed) mutant D12A/D66A, which lacks exo activity due to the substitution of the two catalytic aspartic residues at the exo active site (Soengas *et al*, 1992), and the transfer-deficient mutant N62D (td), in

*Corresponding author. Department of Physics, QB3, Howard Hughes Medical Institute, University of California Berkeley, 608 Stanley Hall, Berkeley, CA 94720-3220, USA. Tel.: +510 643 9706; Fax: +510 643-4500; E-mail: carlos@alice.berkeley.edu
[7]Present address: Departamento Estructura de Macromoléculas, Centro Nacional de Biotecnología (CSIC), Madrid, Madrid 28049, Spain
[8]Present address: Department of Physics, University of Illinois, Urbana, IL 61801, USA
[9]Present address: SIG LLP. Bala Cynwyd, PA 19004, USA

Received: 28 April 2009; accepted: 6 July 2009; published online: 6 August 2009

which the primer transfer reaction is severely reduced due to the poor binding and stabilization of the primer at the exo site, decreasing significantly its degradation (De Vega *et al*, 1996).

The ability to follow in real time the activity of a single polymerase molecule reveals the fluctuations, detected as pauses, of its replication cycle as the protein moves processively along the DNA. The effect of template tension on the pol and exo activities, and on the pausing behaviour of the three polymerases under study, combined with their known biochemical properties, makes it possible to postulate a minimal pathway for the primer transfer process through two obligate intermediate states. As we will show, the proposed kinetic rates, equilibrium constants and conformational changes for each polymerase along the force-induced transfer pathway are compatible with those detected in kinetic and structural studies after the incorporation of an error and during the intramolecular transfer reaction (Donlin *et al*, 1991; Soengas *et al*, 1992; De Vega *et al*, 1996; Shamoo and Steitz, 1999; Hogg *et al*, 2004; Johnson and Beese, 2004; Wang *et al*, 2004; Berman *et al*, 2007). This compatibility, in addition to the known behaviour of DNA under tension (Smith *et al*, 1996), strongly supports the conclusion that mechanical tension applied to a single polymerase–DNA complex promotes the physiological transfer pathway between the polymerase active sites.

Results

Detection of processive single-molecule replication events

To investigate replication by individual Φ29 DNA polymerases, we used optical tweezers to apply mechanical tension between two beads attached to the ends of a 8-kb double-stranded DNA (dsDNA) molecule with a ~400 nucleotide single-stranded gap in the middle (Figure 1A) and monitor at constant force the end-to-end distance change of the DNA (Δx) as the single-stranded template is replicated to dsDNA by Φ29 DNA polymerase (Figure 1B). The number of nucleotides incorporated as a function of time was obtained by dividing the observed distance change (Δx) by the expected change in extension at a given force accompanying the conversion of one single-stranded nucleotide into its double-stranded counterpart (Supplementary Figure 1).

Interestingly, pause events are observed for all three polymerases. Figure 1C shows several pause events along individual replication traces of the ed mutant. These pauses do not represent the exchange of different polymerases at the 3′-end as varying the protein concentration 100-fold (0.2–20 nM) did not change neither the average length of the pauses nor the final length of the primer strand (Supplementary Figure 2). The pause events represent, therefore, protein–DNA complexes in a transient non-active conformation.

Mechanical tension induces the intramolecular pol/exo primer transfer reaction

Figure 2A shows that at the lowest template tensions, the average velocity of the three proteins under study is in agreement with their reported average velocity in bulk (40–60 nt/s) (Soengas *et al*, 1992, 1995; De Vega *et al*, 1996). Interestingly, the replication rate remains constant for a wide range of tensions, suggesting that in this range of forces, mechanical tension does not act directly on the rate-limiting step of the pol cycle of these enzymes (Supplementary Figure 5A). However, template tensions above this range caused the replication rate to decrease rapidly until polymerization is stalled at ~37 pN for the wt and td polymerases and at ~30 pN for the ed mutant (Figure 2A and Supplementary Figure 3). After stalling, polymerization restarted when tension was lowered (Supplementary Figure 3), indicating that no irreversible conformational changes and/or polymerase detachment are induced at these tensions.

At tensions above 46 ± 2 pN, a force-induced exo activity (FIEA) was observed for the wt polymerase even in the presence of dNTPs (Figure 2). A similar effect was reported by previous single-molecule mechanical studies on T7 DNA polymerase (Wuite *et al*, 2000). The rate and length of the FIEA depend on tension but in contrast with the T7 results, are independent of protein concentration (Supplementary Figure 4), indicating that under application of tension the primer transfer process was intramolecular and that the same polymerase molecule degrades processively the primer strand. The processivity of the FIEA (~200 nt at 58 pN, Supplementary Figure 4C) indicates that tension not only

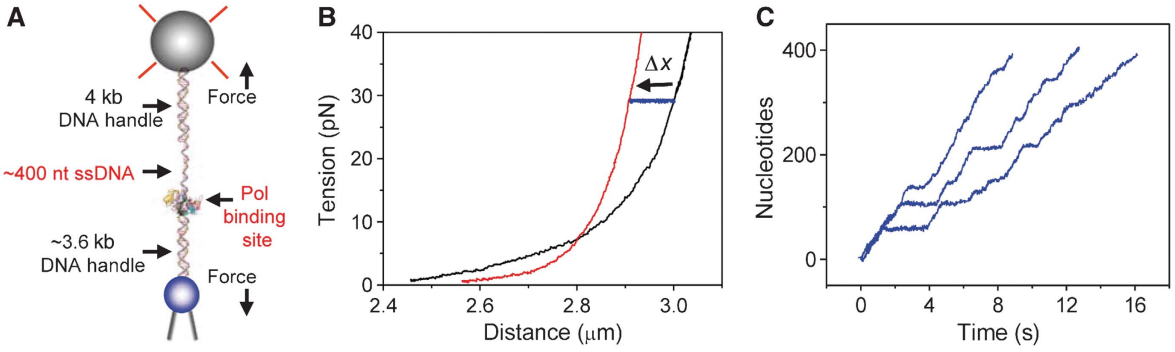

Figure 1 Experimental set-up and detection of single-molecule polymerization events. (**A**) Schematic representation of the experimental set-up (not to scale). A single DNA molecule was tethered to functionalized beads using biotin and digoxigenin moieties at the distal ends of the molecule. One bead (blue) is held in place at the end of a micropipette and the other (grey) by the optical trap. The polymerase is loaded at the 3′ end of the primer–template DNA. (**B**) Replication experiment (29 ± 0.8 pN, wt polymerase) showing the force-extension curves of the initial (black) and final (red) DNA molecules. At constant force, replication shortens the distance (Δx, blue) between the beads. (**C**) Representative replication traces from three independent experiments (22 ± 0.8 pN, ed mutant).

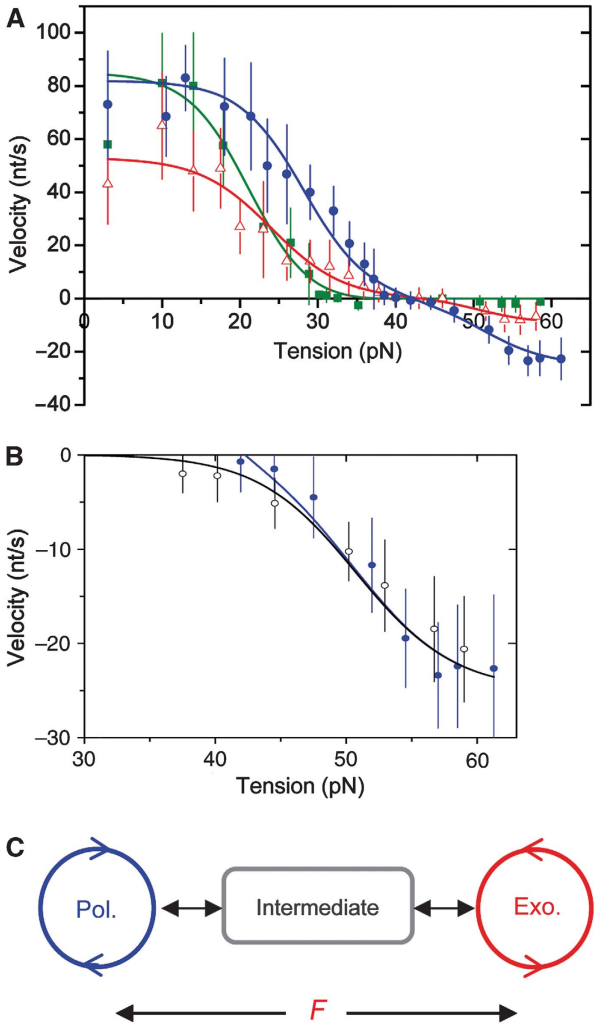

Figure 2 Effect of template tension over the pol and exo rates. (**A**) Average replication rate versus increasing template tension for the wt ($N=125$, blue circles), ed ($N=105$, green squares) and td ($N=102$, red triangles) polymerases. Negative velocities indicate exo activity. Error bars show the standard deviation (s.d.). Solid lines represent the fit of the data with our model (R^2: wt = 0.97, ed = 0.94 and td = 0.93). (**B**) Average exo rate for the wt polymerase versus tension, with 50 μM dNTPs ($N=80$, blue circles) and without dNTPs ($N=30$, black open circles) in the reaction buffer. Error shows s.d. Solid lines are the fit to the data obtained with our model for dNTP ($R^2=0.97$) and no dNTP ($R^2=0.96$) conditions. (**C**) Diagram showing the required intermediate state along the tension coordinate (F) between the active cycles of the polymerase.

triggers the transfer reaction but also stabilizes the editing conformation of the protein. Decreasing the tension below ~37 pN caused processive exonucleolysis to halt and polymerization to resume within a few seconds, 3.2 ± 1.2 s (72%, $N=40$).

To establish the minimum tension that triggers the transfer and stabilization of the primer at the exo domain, we determined the FIEA in the absence of dNTPs to avoid a competent polymerization reaction. Under these conditions, processive (⩾10 nt) exo events were already detected at forces above 30 pN for the wt polymerase (Figure 2B and Supplementary Figure 4B). We note that in the absence of force and dNTPs, Φ29 DNA polymerase presents a spontaneous, non-processive exo activity over a correctly paired

Table I Rates and equilibrium constants of the transfer reaction along the tension coordinate

	wt	ed	td
k_{pol1}	85	80	60
k_{pol2} (0)/d_2	15/0.15	15/0.15	15/0.15
k_{exo}	25	0	8
K_{12} (0)/d_{12}	10^{-3}/1	5×10^{-3}/1	4×10^{-3}/1
k_{12} (0)/d_{12^*}	1.2×10^{-3}/0.95	9×10^{-3}/0.95	10^{-2}/0.95
k_{-12} (0)/d_{-12^*}	1.5/0.05	2.5/0.05	2/0.05
K_p	0.34	0.5	0.63
k_p	0.17	0.20	0.35
k_{-p}	0.5	0.4	0.55
K_t (0)/d_t	10^{-5}/1.1	10^{-4}/1.1	5×10^{-6}/1.1

Values correspond to conditions in the absence of tensions and in the presence of dNTPs (50 μM). Rates are in nt/s and s^{-1}, Arrhenius force dependences (d) in nm. For all parameters, the s.e. is within the ~10% of the showed values.

DNA, which is as slow as 0.008 nt/s (De Vega *et al*, 1999). In our experiments, this activity may occur at low tensions (⩽30 pN), but it would change the distance between the beads at a rate of ~0.09 nm/min (at 20 pN) being its detection obscured by our experimental and thermal drift, 1–10 nm/min (Smith *et al*, 2003). The application of tension brings the otherwise short and slow exo activity, because of the thermodynamically favoured binding of the primer to the pol domain, to an accessible time window.

The observed average degradation rate of the primer for the wt polymerase at saturation, ~25 nt/s, contrasts with its overall exo rate over ssDNA, 500 nt/s (Esteban *et al*, 1994). The difference between these rates suggests that during processive editing one or more rate-limiting steps are required to prepare the trimmed primer for the next excision reaction. In fact, our data shows that individual exo traces present pauses of different durations (Supplementary Figure 4A). These pause events indicate that at least one non-functional state is visited during the processive exo reaction decreasing the average exo degradation. Determination of the nature of these exo-related pause events was impeded by the high experimental noise at high tensions that prevented a detailed analysis of these data. Importantly, the FIEA rates observed for the polymerase mutants (Figure 2A and Table I) reflect their defective exo activity observed in bulk studies after the incorporation of an error (Soengas *et al*, 1992; De Vega *et al*, 1996). This observation supports the interpretation that the pathway followed by the primer leads to the same functional editing conformation when the transfer is triggered by the incorporation of a wrong nucleotide or by the application of tension.

The primer transfer reaction occurs through an obligate intermediate state

Interestingly, the experimental force–velocity curve (Figure 2A) shows a consistent plateau, in which the polymerase is stalled between the pol and exo cycles (from ~37 to 46 pN). If the two cycles were directly connected, their interconversion would be described by a tension-dependent equilibrium constant, $K(f)$ given by (Tinoco and Bustamante, 2002; Bustamante *et al*, 2004):

$$K(f) = K(0)\times \exp(f\times d/k_BT), \quad (1)$$

where f is the applied tension, $K(0)$ the equilibrium constant at zero tension, k_B is the Boltzmann constant, T the absolute temperature and d the distance along the mechanical component of the reaction coordinate separating the two states. This simple two-state model predicts that increased tensions on the template would produce an exponential decrease in polymerization simultaneously with an exponential increase in exonucleolysis (Supplementary data and Supplementary Figure 5A). However, the presence of the plateau between these two activities indicates that tension is not acting directly on the equilibrium between them and at least one non-moving or low-activity intermediate state that accumulates with tension and stalls the polymerase is required between the synthetic and degradative cycles of the polymerase to explain the data (Figure 2C).

Pauses are not responsible for polymerase stalling

Insight into the nature of this intermediate state may be gained from an investigation of the replication pauses (Figure 1C), which may accumulate with force until they finally stall the polymerase. At any given tension, the replication velocity histogram typically shows a bimodal distribution: one peak representing the paused state (centred at 0.2 ± 3 nt/s) and the other the moving state of the protein (Figure 3A). If pauses are responsible for stalling, the ratio between the occupancies of these two states (paused/moving), or the relative occupancy, should be very high at tensions close to 37 pN.

However, Figure 3B shows that for each protein the relative pause occupancy saturates at low values, indicating that even at saturation, the polymerase spends only ~20 to 40% of its total time paused. Therefore, the occupancy of the pause state cannot be responsible for tension-induced stalling characterized here. This unexpected behaviour was confirmed by an independent method for pause scoring based on the statistical analysis of velocity fluctuations (Supplementary Figure 6). The obtained low pause frequency and short pause length (Figure 4) at 37 pN for the wt and td polymerases also support this observation. For the ed mutant, the particular conditions that lead to the sudden increase in relative pause occupancy and subsequent stalling at ~30 pN are discussed below.

The initial pol cycle and the pause state are connected through an intermediary

The relative occupancy described above is a direct measurement of the effective 'equilibrium constant' for the interconversion between the pol cycle and the pause state (Supplementary data). Therefore, if the paused state is accessed from the moving state as a direct consequence of the tension applied to the template, the tension dependence of the relative pause occupancy would follow an expression similar to equation (1) and we would expect it to increase exponentially with force (Supplementary Figure 5B).

Instead, Figure 3B shows an early saturation of the relative pause occupancy with tension. Deviation from a single exponential indicates that the effect of tension on the interconversion between the moving and paused states is not direct. In fact, this saturation argues for the existence of an additional intermediate located between the initial pol cycle (pol1) and the pause state of the protein (Figure 3C). The initial sharp increase in the relative pause occupancy indicates that access to this intermediate from the initial pol1 is

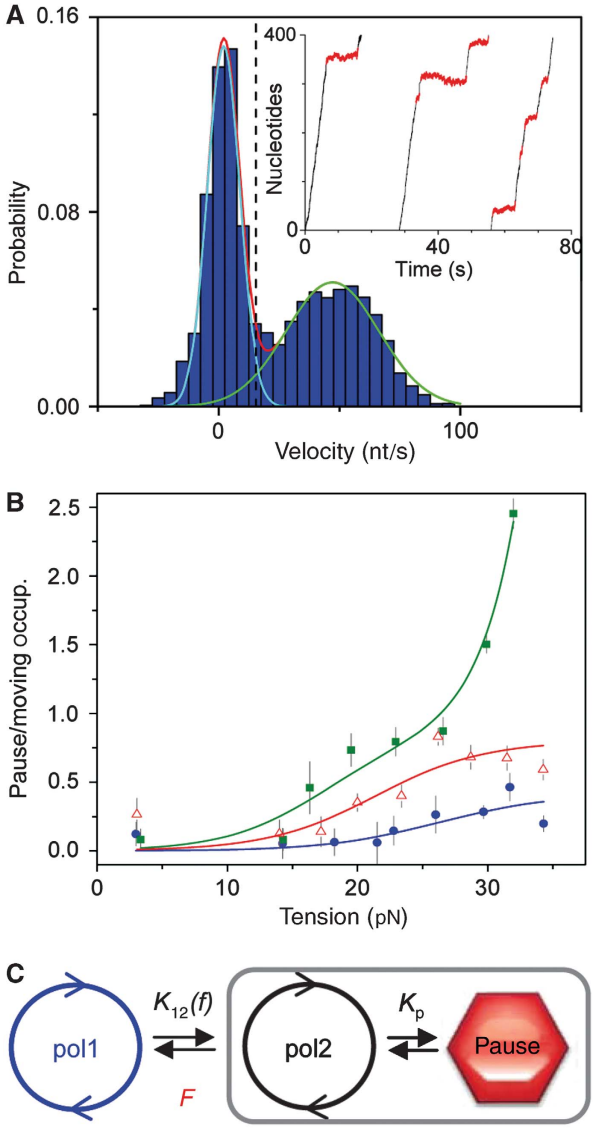

Figure 3 Pause identification and tension effect on the relative pause occupancy. (**A**) Velocity distribution of five independent runs (23 ± 0.8 pN, ed polymerase) fit by the sum of two Gaussians (red line). Velocity points below the threshold, velocity value of two standard deviations below the moving state peak (dashed black line), are considered as pause events along the replication run (in red, inset), resolution of 0.5–1 s. (**B**) The effect of tension on the relative pause occupancy for the wt (blue dots), ed (green squares) and td (red triangles) polymerases. The relative pause occupancy was determined from the ratio of the areas under each peak of the velocity distribution shown in panel A. The relative pause occupancy saturates above 25 pN for the wt and td polymerases and between 19 and 27 pN for the ed mutant. Values are displayed as mean ± standard error (s.e.). Solid lines are the fit of the data with our model. R^2: wt = 0.66, ed = 0.97 and td = 0.75. (**C**) Diagram showing the access from the initial replication cycle (pol1) to the pause state (in red) through a functional conformational intermediate (pol2). F denotes the tension sensitive transition.

force-sensitive and the ensuing saturation requires that the equilibrium between the intermediate and the paused state, K_p, be insensitive to the template tension. Importantly, this new intermediate must be a moving or polymerization-competent cycle (that we call pol2), as direct access to a non-active state in a tension-sensitive manner would lead to a continuous exponential increase in the relative pause occupancy, which is not observed.

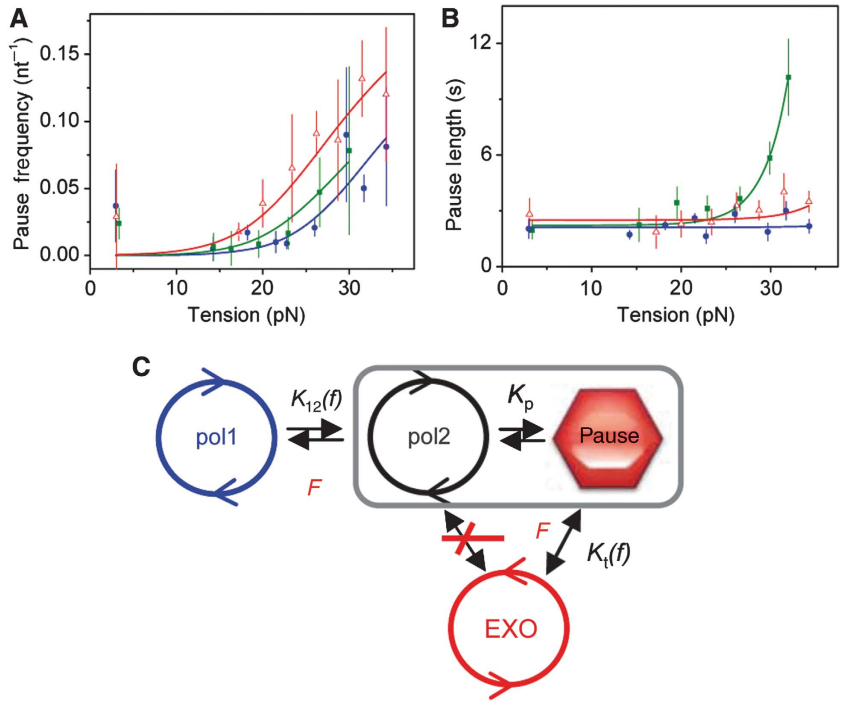

Figure 4 Pause state features. For all figures wt (blue), td (red) and ed (green). (**A**) Pause frequency (defined as the inverse of the distance between pauses, pauses/100 nt) versus template tension. Error bars show the s.d. Solid lines are the fit to the data (R^2: wt = 0.58, ed = 0.98, td = 0.91) with the pause frequency expression derived from our model (Supplementary data). (**B**) Average pause length (defined as the duration of the pause in seconds) versus tension. Error bars show s.e. Solid lines, data fit with our current model (R^2: wt = 0.5, ed = 0.93, td = 0.45). (**C**) Diagram showing the effect of template tension (F) shifting the equilibrium from the pol to exo cycle through the pause state.

Therefore, the relative occupancy of the pause-moving state equilibrium is given by (see Supplementary data for additional P_{oc} definitions):

$$P_{oc} = \frac{K_{12}(f)}{1 + K_{12}(f)} K_p, \qquad (2)$$

that saturates at the value of K_p, and $K_{12}(f)$ is the tension-dependent equilibrium constant of the pol1 to pol2 transition (Table I). An Arrhenius force dependence for $K_{12}(f)$ with a distance d_{12} of 1 ± 0.1 nm was required to fit accurately the sharp initial exponential increase in the relative occupancy observed for all three polymerases, indicating that a conformational change of at least this magnitude along the tension coordinate occurs in the polymerization complex from the pol1 to the pol2 states. As discussed below, this conformational change is compatible with the release of the template strand initially organized inside the polymerase.

Pol2 cycle is responsible for the wt and td polymerases stalling

Our data show that neither the pol1 replication rate, k_{pol1} (Supplementary Figure 5), nor the relative pause occupancy (Figure 3B) is responsible for the tension-induced stalling of the wt and td polymerases. Therefore, the pol2 replication rate (k_{pol2}) should depend on force to explain the observed tension-induced replication stalling.

This model was tested by analysing the force–velocity behaviour of the polymerase (Figure 2A). As expected from the above arguments, fit to the data (see velocity definition in the next section and Supplementary data) requires a pol2 replication rate, $k_{pol2}(0) = 15 \pm 5$ nt/s, with a force-dependent associated distance of $d_2 = 0.15 \pm 0.05$ nm for all three proteins. We note that the observed sharp decrease in the average pol rate with force that eventually leads to stalling is mainly due to the strong effect of tension ($d_{12} \sim 1$ nm) in shifting the equilibrium ($K_{12}(f)$) from the 'fast' pol1 to the 'slow' pol2 conformation, which presents a rate of $k_{pol2}(f) \sim 1 \pm 3$ nt/s at 37 pN (Supplementary data and Supplementary Figure 7). The particular conditions that lead to the stalling of the ed mutant are discussed below.

The pause state is an obligate intermediate between the pol and exo cycles

Once the polymerase is stalled, the transfer of the primer to the exo cycle could, in principle, occur either directly from the pol2 cycle or from the paused state (Figure 4C). We found that our data are only compatible with a model in which the pause acts as the only obligatory intermediate between the pol and exo cycles (Supplementary data and Supplementary Figure 5).

When the pause acts as an intermediate, access to the FIEA (with and without dNTPs, Figure 2B) is determined by a velocity, v, given by

$$v = \frac{k_{pol1} + K_{12}(f) \times (k_{pol2}(f) - k_{exo} \times K_p \times K_t(f))}{1 + K_{12}(f) \times (1 + K_p + K_p \times K_t(f))}, \qquad (3)$$

where k_{exo} is the force-induced exo rate at saturation (~ 25 nt/s), and all terms except for $K_t(f)$, the force-dependent equilibrium constant of the transfer reaction from the paused state to the exo cycle, are determined by previous fits (Table I). The value of K_t obtained for the Φ29 DNA polymerase, 10^{-5}, contrasts with the ~ 10 times higher $K_t(10^{-4})$ reported for the T7 DNA polymerase (Donlin *et al*, 1991). This difference could be compensated in terms of fidelity by

the high nucleotide binding discrimination shown by the Φ29 DNA polymerase, which is exactly ~10 times higher than T7 DNA polymerase (Saturno *et al*, 1995). The smaller $K_t(f)$ value for the td mutant (Table I) is in accordance with the poor stabilization of the primer at the exo site described for this polymerase (De Vega *et al*, 1996).

Importantly, fit to the fast increase in FIEA observed for the wt and td polymerases requires a conformational change of $d_t = 1.1 \pm 0.1$ nm along the tension coordinate, which is compatible with the fraying of the primer–template DNA structure that is needed for the transfer reaction (discussed below) (Shamoo and Steitz, 1999; Hogg *et al*, 2004).

Destabilization of the primer–template structure triggers the transfer and stalls the advance of the ed mutant

Recent single-molecule studies showed that longitudinal tensions on DNA above ~30 pN start to unwind the dsDNA helix (Gore *et al*, 2006). Interestingly, ~30 pN is the tension value that induces the transfer and stabilization of the primer at the exo domain for the wt polymerase detected in the absence of dNTPs (Figure 2B). These observations suggest that tension-induced dsDNA unwinding could promote, in the presence of the protein, fraying of the primer–template structure and the concomitants intramolecular primer transfer and exo reactions.

For the ed mutant tensions within this range (~30 pN) cause a sharp increase in pause length (Figure 4B) (and therefore, in pause occupancy), which is sufficient to stall the advance of this polymerase. The main difference between the ed mutant and the two other proteins is that the former has no exo activity. Therefore, tensions that trigger the transfer process would stall this polymerase by shifting the primer to and stabilizing it at the inactive exo domain, which is scored as an additional pause event in our experiments.

According to our model, the pause length definition is given by (Supplementary data)

$$P_L = \frac{1}{k_{-p}}(1 + K_t(f)), \tag{4}$$

where k_{-p} is the exit rate from the pause state. Fit to the sharp increase in pause length observed for the ed mutant (Figure 4B), which reflects the transfer reaction, requires exactly a force dependence of $d_t = 1.1 \pm 0.1$ nm, indicating that the same conformational change is induced by tension in all three proteins.

It can be noted that this expression fits accurately the data for all three proteins even when all the values are constrained by previous fits, underlying the robustness of our model. Moreover, the value of $k_{-p} \sim 0.5\ s^{-1}$ (wt polymerase) is in keeping with the measured time delay to recover polymerization after exonucleolysis when tension is lowered, 3.2 ± 1.2 s ($N = 29$), supporting further the idea of the pause state as an obligatory intermediate between the pol and exo cycles.

Discussion

The strong binding of the template and primer–template at the pol domain favours the polymerization active conformation

According to our model, tension does not affect directly the replication cycle, instead shifts the equilibrium from the initial pol1 cycle to the new, tension-sensitive, 'slow' pol2 cycle, which stalls at ~37 pN. Transition between these two cycles implies a conformational change of $d_{12} = 1 \pm 0.1$ nm, a distance compatible with the gain of ~3 ssDNA nucleotides along the mechanical coordinate (Supplementary data), which may originate from the release of the ~3 template nucleotides initially organized inside the protein (Berman *et al*, 2007) (Figure 5B). After this interpretation, the low equilibrium constant of the pol1 to pol2 transition, K_{12} *(f)*

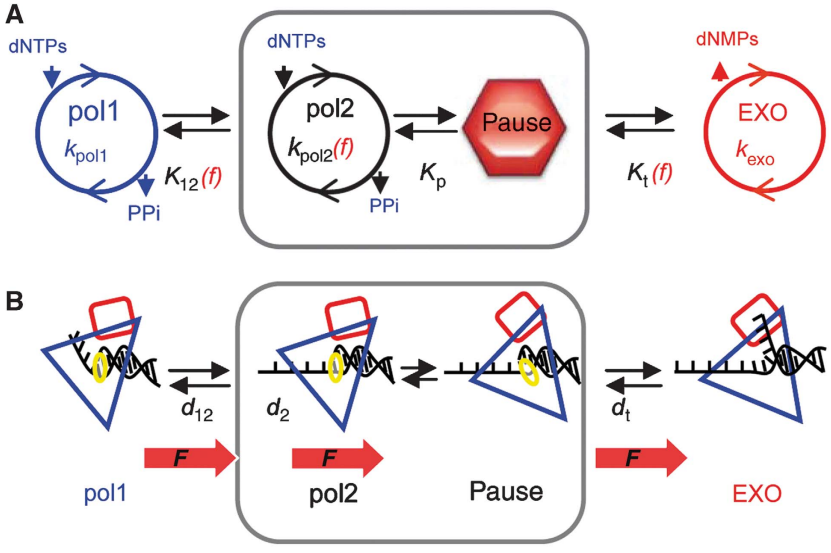

Figure 5 Conformational dynamics along the tension coordinate. (**A**) Diagram showing the proposed transitions from the pol to the exo cycles. (**B**) Schematic representation of the proposed conformational transitions: pol1, initial protein–DNA conformation with DNA bound at the pol domain (blue triangle) and the DNA correctly aligned with the pol active site (yellow circle). pol2, polymerase loses contact with template strand, the duplex DNA is still correctly aligned with the pol active site but translocation is affected (d_2). Under these conditions, the protein could explore alternative non-functional conformations (pause state) where the DNA and the pol active site may not be correctly aligned. From the pause state, the primer is transferred to the exo domain (red square) by the unwinding of 4–5 nucleotides of the dsDNA. Red arrows show the localization of the tension sensitive transitions.

(Table I), can be interpreted as a measure of the strong affinity of the polymerase for the template strand (Supplementary data).

Our results bring out the importance of the polymerase–DNA interactions to modulate the polymerization cycle. On the one hand, template interactions are crucial to modulate the replication velocity. The proposed pol2 conformation represents a DNA–polymerase complex with loose template interactions, which can explain its decreased rate (k_{pol2} (0) $\sim$15 nt/s) and tension sensitivity, d_2. Interestingly, the distance to the transition state for stalling of the pol2 cycle, d_2 $\sim$0.15 nm, is in excellent agreement with the $\sim$0.15 nm displacement the templating nucleotide undergoes during each replication step (Berman *et al*, 2007), suggesting that tension prevents the shift of this strand required for proper template reading and nucleotide incorporation.

On the other hand, it is remarkable that even under conditions that promote template loosening, the polymerase remains active (pol2 cycle), pointing out the strong affinity of the pol active site for the correctly paired primer–template DNA structure. This could be crucial for the proper coordination between the synthetic and degradative activities of the protein, as it favours a replication-competent conformation even under conditions that may transiently slow down the replication advance, avoiding unnecessary excursions to the exo domain.

The pause state could work as a fidelity checkpoint

As replication velocity decreases, pauses begin to accumulate. The independence on tension of the moving–pause state equilibrium indicates that any conformational change associated with this step is either independent of template tension or orthogonal to the direction of applied tension. Therefore, no conformational change or strand opening is expected along the dsDNA structure (which is in the direction of pulling) when the polymerase visits the pause state (Figure 5B).

As the pause state constitutes a direct intermediate between the synthetic and degradative cycles, it may have a crucial role in the coordination of these opposite activities. At this state, the polymerase may check the integrity of the DNA structure before adopting the final degradative (exo) conformation, acting in this way, as an off-pathway fidelity checkpoint favouring selective degradation of insertion errors. In addition to DNA lesions, mismatches or mechanical tension, other factors known to modify the template elastic properties and its local structure, such as protein binding or unbinding events (Soengas *et al*, 1995), and ionic conditions (Supplementary Figure 1) could affect the polymerase–template interactions and shift the equilibrium towards the pol2 pause-prone state during a processive replication event. In fact, we observed pause events that occur spontaneously at low template tensions (Figure 4A). In this context, checking of the primer–template integrity would favour selective degradation of insertion errors: after template release, if the dsDNA is correctly paired, visiting the pause state will allow the DNA to recover its functional binding to the pol site and replication will resume avoiding unnecessary proofreading of the primer. If, instead, the dsDNA structure is disrupted, the instability of the template–primer structure and the affinity of the exo domain for the ssDNA would rapidly shift the equilibrium towards the exo cycle (see below).

Primer–template conformation is determinant to trigger a full transfer reaction

Our model shows that once the polymerase is in the inactive pause state, visit to and stabilization at the final exo configuration would be an improbable event as the equilibrium for this reaction is strongly shifted towards the pause ($K_t(f)$ $\sim 10^{-5}$). Therefore, following with the above discussion, decreasing the affinity of the pol active site for the template strand could initiate the transfer reaction but it is not sufficient to trigger the full process. Disruption of the proper dsDNA primer–template structure should be necessary to induce a successful transfer process. As reported earlier for Families A and B DNA polymerases, any condition that stimulates strand separation or instability of the local primer–template terminus, as a temperature or sequence context, affects the proofreading activity (Bessman and Reha-Krantz, 1977; Petruska and Goodman, 1985; Bloom *et al*, 1994; Carver *et al*, 1994; Marquez and Reha-Krantz, 1996; Hariharan and Reha-Krantz, 2005). In our experiments, tensions above 30 pN disrupt or unwind the dsDNA sufficiently to promote, in the presence of the protein, the transfer reaction, which in turn induces a net exo activity that can be readily detected at these forces in the absence of dNTPs (Figure 2B and Supplementary Figure 4A). Interestingly, the transfer equilibrium constant, $K_t(f)$, presents a strong force dependence, $d_t = 1.1 \pm 0.1$ nm (for the three polymerases under study), indicating that fraying of $\sim$4 to 5 bp of dsDNA is required for the single-stranded primer to reach the exo active site (Figure 5B and Supplementary data).

We note that the FIEA corresponds to processive primer degradation events (Supplementary Figure 4) and, therefore, to the mechanical stabilization of the polymerase in the functional editing conformation. An important consequence of this interpretation is that the onset of saturation in the FIEA degradation rate should be independent of the dNTP concentration, as shown experimentally and also reflected by our model (Figure 2B).

Tension-induced distortions of DNA mimic the physiological transfer process

Several lines of evidences supported by structural and kinetic data on Φ29 and related DNA polymerases show that mechanical tension is a valid variable to study the dynamics of the physiological proofreading reaction.

Structural studies using the *Bacillus* DNA polymerase I fragment showed that 6 of all 12 possible mismatches induce long-range DNA structural distortions of the DNA helix affecting the primer–template and the template strand interactions at the pol active site, the latter appearing dislodged from the pol site (Johnson and Beese, 2004). In a similar way to the incorporation of these types of errors, we have proved that tension induces the replication stalling and the primer transfer reactions by disrupting specifically the template and primer–template interactions with the pol domain, although the final structure of the polymerase–DNA complex in each situation probably differs.

The tension-induced structural modifications of the template determined by our model are compatible with recent structural studies showing that the template length organized inside this protein is three nucleotides ($d_{12} \sim 1$ nm) and the shift of the templating nucleotide predicted for each translocation step is 0.15 nm ($d_2 \sim 0.15$ nm) (Berman *et al*, 2007).

After template release, entering the intermediate pause state does not involve DNA opening (K_p is force independent), which may resemble the structure of pol–exo intermediate states observed for the related RB69 DNA polymerase in which the template and a still paired primer–template structure are displaced from the pol active site towards the exo domain (Hogg *et al*, 2004). Finally, the tension induced fraying of the primer–template structure required to reach the exo active site and stabilize the editing conformation of the protein, 4–5 nucleotides ($d_t \sim 1.1$ nm) is in keeping with the 3–4 nucleotides frayed at the exo site in the editing conformation of the related RB69 DNA polymerase (Shamoo and Steitz, 1999; Hogg *et al*, 2004).

Moreover, the exo rates measured under application of tension for the ed and td mutants agree very well with their decreased rate observed in biochemical bulk experiments in the absence of force, suggesting that the same functional editing configurations is adopted in both conditions. We note that all DNA polymerases with 3′–5′ exonuclease activity and with known structures have a cluster of highly conserved side chains at the exo active site, including the ligands for the two divalent metal ions required for catalysis (i.e., the aspartic acids D12 and D66 of the Phi29 polymerase) and the residues involved in the stable transfer of the primer and interaction with the ssDNA (i.e., the lysine residue 66 of the Phi29 polymerase). These residues are actively involved in the transfer, stabilization and degradation of the primer at the exo active site after the incorporation of an error (for a compilation of conserved exo domain sequences, see Bernad *et al*, 1989; Blanco *et al*, 1991; Morrison *et al*, 1991; Blanco *et al*, 1992; de Vega *et al*, 1998). Therefore, the results with the exo-deficient D12A/D66A (ed) and transfer-deficient N62D (td) mutants are of particular significance proving that the same residues involved in the transfer and stabilization of the primer in physiological conditions are also involved in the transfer reaction under application of tensions and, therefore, the same transfer pathway is used in both conditions.

We have shown that the application of increasing mechanical tensions to single polymerase–DNA complexes furnishes a new, controlled, quantitative approach to determine intermediates of the proofreading reaction and to measure directly the kinetic rates, equilibrium constants and conformational changes associated with their interconversion. This method opens a new way to study the response of proofreading dynamics to DNA sequence and dNTP concentration effects.

Materials and methods

DNA substrate

The *pBacgus11* (Novagen) dsDNA vector (8041 bp, 0.66 μg) was nicked at a single position with the *NBbvIA* nicking enzyme (New England Biolabs) following the manufacturer's reaction conditions. The newly generated 3′ end was cleaved with 0.3 μl of *Exonuclease III* (100 U/μl) (New England Biolabs) in the nicking enzyme buffer for 2 min at 37°C, in order to create a ~400 nucleotide long ssDNA gap. The reaction was stopped with EDTA (10 mM) and Proteinase K (0.1 μg/μl) for 1 h at 37°C. The DNA was purified with the Qiagen PCR Purification Kit (Qiagen) and digested with *Hin*dIII and *Bam*HI restriction endonucleases (New England Biolabs) following the manufacturer's reaction conditions. The length of the single stranded gap was monitored by denaturing agarose gel electrophoresis.

Digoxigenin (Dig)-labelled DNA handles were generated by PCR amplification of the multiple cloning site of the pUC19 DNA vector (New England Biolabs) using 0.2 mM of dATP, dCTP and dGTP, 0.13 mM of dTTP and 0.07 mM of DIG-dUTP. PCR products were digested and gel-purified to recover a 3650-bp DIG-modified *Bam*HI. The biotinylated DNA handle consists of a 3′ biotin-labeled DNA oligonucleotide with the 5′ end complementary to the *Hin*dIII sequence (5′-AGCTATCAGGCTACAAA-3′). The handles were ligated using the T4 DNA ligase (New England Biolabs) to their complementary positions at the hybrid DNA molecule following the manufacturer's reaction conditions.

Optical tweezers experiments

We used a dual-beam optical tweezers instrument (Smith *et al*, 2003) to manipulate individual DNA molecules. Φ29 DNA polymerase (0.2, 2 and 20 nM) was flowed into the chamber diluted in the reaction buffer (50 mM Tris–HCl, 20 mM ammonium sulphate, 1 mM DTT, 10 mM $MgCl_2$, 4% glycerol (w/v), 0.05% Tween 20 (w/v), 0.1 μg/ml BSA and the four dNTPs (50 μM)). Generation and purification of the wt, ed and td polymerases was described elsewhere (Soengas *et al*, 1992; De Vega *et al*, 1996, 1999).

Data were collected at 60 Hz at 22 ± 1°C and measurements were carried out in two modes: 'constant force feedback', in which the distance between the beads was adjusted to maintain a constant tension in the DNA, and 'no feed back', in which the distance between the optical trap and the pipette was held constant.

Replication rate determination

The number of replicated nucleotides was calculated as described earlier (Wuite *et al*, 2000), using the worm-like chain model for polymer elasticity with a persistent length of $P = 53$ nm and stretch modulus $S = 1200$ pN/nm for dsDNA, and $P = 0.75$ nm and $S = 800$ pN/nm for the ssDNA template. Below 20 pN, the ssDNA values were obtained from the experimental curves (Supplementary Figure 1).

The average velocity at each tension was calculated from a linear fit to the arrangement in chains of several individual activities taken at constant force feedback within the same range of tensions (± 1–2 pN). Instantaneous replication rates were obtained from a linear fit of the number of replicated nucleotides over a sliding time window of 0.7 s (50 data points) for all velocities. Velocity distributions were determined from the histogram of the instantaneous replication rate using a bin of five nucleotides per second.

Replication velocity at 8–10 pN where the end-to-end extensions between ss and ds nucleotides are very similar was calculated as follows. Replication was initiated holding the force constant at ~12 pN, as soon as replication started the force was lowered, fixed to 8–10 pN for a few seconds and then increased again to 12 pN for polymerization to finish. The average velocity at 8–10 pN was estimated by measuring the time interval and the distance change at 12 pN between the initial and final 'directly measured' replication events.

Pause and moving states occupancy

We defined the occupancy of the paused and moving states of the protein as the fraction of time the polymerase spends in that state. To determine occupancies of the pause and moving states, the velocity distribution was fit to a sum of two Gaussians. The occupancy of each state was calculated as the area under that state divided by the total area of the distribution. The relative occupancy was determined from the ratio of the areas under each peak. The calculation of pause occupancy was also achieved with an independent method based on the statistical analysis of velocity fluctuations (Supplementary Figure 6).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank JR Moffitt for critical reading of the paper, E Galburt and S Grill for help with data analysis and helpful discussions. This work was supported by the National Institutes of Health (CB), the Spanish Ministry of Education and Science and Comunidad de Madrid (MS). YRC acknowledges the Burroughs Wellcome Fund, Career Award at the Scientific Interface. BI and CB designed the

experiments; BI performed the experiments; BI, YRC and SP analysed the data; SBS, JML and MS provided technical and biochemical tools; and BI, YRC and CB wrote the paper.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Benkovic SJ, Valentine AM, Salinas F (2001) Replisome-mediated DNA replication. *Annu Rev Biochem* **70:** 181–208

Berman AJ, Kamtekar S, Goodman JL, Lázaro JM, de Vega M, Blanco L, Salas M, Steitz TA (2007) Structures of Φ29 DNA polymerase complexed with substrate: the mechanism of translocation in B-family polymerases. *EMBO J* **26:** 3493–3505

Bernad A, Blanco L, Lazaro J M, Martin G, Salas M (1989) A conserved 3′–5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases. *Cell* **59:** 219–228

Bessman MJ, Reha-Krantz LJ (1977) Studies on the biochemical basis of spontaneous mutation. V. Effect of temperature on mutation frequency. *J Mol Biol* **116:** 115–123

Blanco L, Bernad A, Lázaro JM, Martín G, Garmendia C, Salas M (1989) Highly efficient DNA synthesis by the phage Φ29 DNA polymerase. *J Biol Chem* **264:** 8935–8940

Blanco L, Bernad A, Blasco MA, Salas M (1991) A general structure for DNA-dependent DNA polymerases. *Gene* **100:** 27–38

Blanco L, Bernad A, Salas M (1992) Evidence favouring the hypothesis of a conserved 3′–5′ exonuclease active site in DNA-dependent DNA polymerases. *Gene* **112:** 139–144

Bloom LB, Otto MR, Eritja R, Reha-Krantz LJ, Goodman MF, Beechem JM (1994) Pre-steady-state kinetic analysis of sequence-dependent nucleotide excision by the 3′-exonuclease activity of bacteriophage T4 DNA polymerase. *Biochemistry* **33:** 7576–7586

Bustamante C, Chemla YR, Forde NR, Izhaky D (2004) Mechanical processes in biochemistry. *Annu Rev Biochem* **73:** 705–748

Carver Jr TE, Hochstrasser RA, Millar DP (1994) Proofreading DNA: recognition of aberrant DNA termini by the Klenow fragment of DNA polymerase I. *Proc Natl Acad Sci USA* **91:** 10670–10674

De Vega M, Lázaro JM, Salas M, Blanco L (1996) Primer-terminus stabilization at the 3′–5′ exonuclease active site of Φ29 DNA polymerase. *EMBO J* **15:** 1182–1192

De Vega M, Lázaro JM, Salas M, Blanco L (1998) Mutational analysis of phi29 DNA polymerase residues acting as ssDNA ligands for 3′–5′ exonucleolysis. *J Mol Biol* **279:** 807–822

De Vega M, Blanco L, Salas M (1999) Processive proofreading and spatial relationship between polymerase and exonuclease active sites of bacteriophage Φ29 DNA polymerase. *J Mol Biol* **292:** 39–51

Donlin MJ, Patel SS, Johnson KA (1991) Kinetic partitioning between the exonuclease and polymerase sites in DNA error correction. *Biochemistry* **30:** 537–546

Esteban JA, Soengas MS, Salas M, Blanco L (1994) 3′–5′ Exonuclease activity site of Φ29 DNA polymerase. *J Biol Chem* **269:** 31946–31954

Fidalgo da Silva E, Reha-Krantz LJ (2007) DNA polymerase proofreading: active site switching catalyzed by the bacteriophage T4 DNA polymerase. *Nucleic Acids Res* **35:** 5452–5463

Freemont PS, Friedman JM, Beese LS, Sanderson MR, Steitz TA (1988) Cocrystal structure of an editing complex of Klenow fragment with DNA. *Proc Natl Acad Sci USA* **85:** 8924–8928

Gore J, Bryant Z, Nöllmann M, Le MU, Cozzarelli NR, Bustamante C (2006) DNA overwinds when stretched. *Nature* **441:** 836–839

Hariharan C, Reha-Krantz LJ (2005) Using 2-aminopurine fluorescence to detect bacteriophage T4 DNA polymerase-DNA complexes that are important for primer extension and proofreading reactions. *Biochemistry* **44:** 15674–15684

Hogg M, Wallace SS, Doublie S (2004) Crystallographic snapshots of a replicative DNA polymerase encountering an abasic site. *EMBO J* **23:** 1483–1493

Hogg M, Aller P, Konigsberg W, Wallace SS, Doublié S (2007) Structural and biochemical investigation of the role in proofreading of a β hairpin loop found in the exonuclease domain of a replicative DNA polymerase of the B family. *J Biol Chem* **282:** 1432–1444

Johnson SJ, Beese LS (2004) Structures of mismatch replication errors observed in a DNA polymerase. *Cell* **116:** 803–816

Kamtekar S, Berman AJ, Wang J, Lázaro JM, de Vega M, Blanco L, Salas M, Steitz TA (2004) Insight into the displacement and processivity from the crystal structure of the protein-primed DNA polymerase of bacteriophage Φ29. *Mol Cell* **16:** 609–618

Kunkel TA, Bebenek K (2000) DNA replication fidelity. *Annu Rev Biochem* **69:** 497–529

Marquez LA, Reha-Krantz LJ (1996) Using 2-aminopurine fluorescence and mutational analysis to demonstrate an active role of bacteriophage T4 DNA polymerase in strand separation required for 3′→5′ exonuclease activity. *J Biol Chem* **271:** 28903–28911

Morrison A, Bell JB, Kunkel TA, Sugino A (1991) Eukaryotic DNA polymerase amino acid sequence required for 3′–5′ exonuclease activity. *Proc Natl Acad Sci USA* **88:** 9473–9477

Petruska J, Goodman MF (1985) Influence of neighbouring bases on DNA polymerase insertion and proofreading fidelity. *J Biol Chem* **260:** 7533–7539

Rodriguez I, Lázaro JM, Blanco L, Kamtekar S, Berman AJ, Wang J, Steitz TA, Salas M, de Vega M (2005) A specific subdomain in Φ29 DNA polymerase confers both processivity and strand-displacement capacity. *Proc Natl Acad Sci USA* **102:** 6407–6412

Saturno J, Blanco L, Salas M, Esteban JA (1995) A novel kinetic analysis to calculate nucleotide affinity of proofreading DNA polymerases. *J Biol Chem* **270:** 31235–31243

Shamoo Y, Steitz TA (1999) Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* **99:** 155–166

Smith SB, Cui Y, Bustamante C (1996) Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **271:** 795–799

Smith SB, Cui Y, Bustamante C (2003) Optical-trap force transducer that operates by direct measurement of light momentum. *Methods Enzymol* **361:** 134–162

Soengas MS, Esteban JA, Lázaro JM, Bernad A, Blasco MA, Salas M, Blanco L (1992) Site directed mutagenesis at the Exo III motif of Φ29 DNA polymerase. Overlapping structural domains for the 3′–5′ exonuclease and strand-displacement activities. *EMBO J* **11:** 4227–4237

Soengas MS, Gutierrez C, Salas M (1995) Helix-destabilizing activity of Φ29 single-stranded DNA binding protein: effect on the elongation rate during strand displacement DNA replication. *J Mol Biol* **253:** 517–529

Strick R, Knopf CW (1998) DNA binding properties and processive proofreading of herpes simplex virus type 1 DNA polymerase. *Biochim Biophys Acta* **1388:** 315–324

Subuddhi U, Hogg M, Reha-Krantz LJ (2008) Use of 2-aminopurine fluorescence to study the role of the β hairpin in the proofreading pathway catalyzed by the phage T4 and RB69 DNA polymerases. *Biochemistry* **47:** 6130–6137

Tinoco I, Bustamante C (2002) The effect of force on thermodynamics and kinetics of single molecule reactions. *Biophys Chem* **101:** 513–533

Wang J, Sattar AK, Wang CC, Karam JD, Konigsberg WH, Steitz TA (1997) Crystal structure of a pol alpha family replication DNA polymerase from Bacteriophage RB69. *Cell* **89:** 1087–1099

Wang CX, Zakharova E, Li J, Joyce CM, Wang J, Konigsberg W (2004) Pre-steady-state kinetics of RB69 DNA polymerase and its exo domain mutants: effect of pH and thiophosphoryl linkages on 3′–5′ exonuclease activity. *Biochem* **43:** 3853–3861

Wuite GJL, Smith SB, Young M, Keller D, Bustamante C (2000) Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature* **404:** 103–106

Yang J, Zhuang Z, Roccasecca RM, Trakselis MA, Benkovic SJ (2004) The dynamic processivity of the T4 DNA polymerase during replication. *Proc Natl Acad Sci USA* **101:** 8289–8294

